DRAFT

JUN 28 1982

ESTIMATION OF STREAMFLOWS

AND

THE REACH FILE

Prepared for:

U. S. ENVIRONMENTAL PROTECTION AGENCY
Monitoring Branch
Washington, D. C. 20460

Project Officer: C. Robert Horn

Prepared by:

W. E. GATES AND ASSOCIATES, INC.

TABLE OF CONTENTS

LIST OF FIGURES

SECTION I

INTRODUCTION

BACKGROUND

The Federal Water Pollution Control Act of 1972 and subsequent amendments thereto established a national objective of restoring and maintaining the quality of the Nation's waters. The U.S. Environmental Protection Agency (USEPA) has responded to this stated objective through the implementation of a variety of planning, regulatory, and enforcement procedures and programs. A significant element of many of USEPA's programs is the identification of the magnitude of water quality problems. This is an extremely important issue in the relatively new field of toxic waste management where USEPA has successfully delineated toxic waste dischargers, and their impact on in situ water quality, biotic communities, as well as sources of water supply.

The organizing framework for delineating, storing, and manipulating data on toxic wastes emanating from industrial and municipal dischargers and nonpoint sources of pollution is USEPA's Reach File. The Reach File is a computerized data base management system which permits the explicit definition of the topologic and geodetic structure of surface waters in the contiguous United States and the storage of a vast array of physical data on each water course. There are currently approximately 60,000 stream segments or reaches represented in the Reach File. The existing data base includes the geodetic coordinates of stream traces, stream names, reach lengths, and other information including data which defines linkages between reaches. Data related to streamflow gages, pollutant sources, and water supply intakes exist in other USEPA data bases which have been indexed to the Reach File.

USEPA is currently developing the technology to route pollutants on a conservative and/or non-conservative basis through the use of the Reach File and to display the results of the routings in a variety of tabular and graphic exhibits. In order to utilize the full capacity of this new technology USEPA required the development of specific flow data for the reaches currently in the Reach File. To this end, W. E. Gates and Associates, Inc. (WEG/A) was selected to provide average and low flow estimates as well as monthly average flows for all reaches in the current File. The objective of this work was to develop the necessary flow data in a totally consistent manner suitable for incorporation into a national data base, utilizing flow-related physical attributes, existing gaging station data and computerized techniques to achieve the required consistency across the Nation.

The remainder of this report contains the description of the methodology employed to attain the stated objective.

OVERVIEW

The results of WEG/A's efforts include a file containing estimates of average annual streamflow, low flow, and average monthly streamflow for each of the structurally connected stream reaches in the Reach File. The methodology developed by WEG/A makes maximum use of information associated with streamflow gaging stations previously associated with specific reaches as well as the ancillary data on the connectivity of the network and stream lengths contained in the File. These data were used in conjunction with defined drainage areas and consistent geomorphological/climatic data to produce a single, uniform procedure for estimating streamflows by reach nation-wide.

The source of the basic stream connectivity information required by the flow estimating technology was the Reach File developed by USEPA for each of the 18 U. S. Geological Survey (USGS) Hydrologic Regions in the contiguous United States. Prior to using the File it was necessary to identify stream connectivity and code/format errors which would generate erroneous streamflow estimates. After USEPA corrected the so-called "fatal" connectivity errors in the Reach File, the File was reformatted into a direct access Analytical Reach File (ARF) in order to support the analysis required for streamflow estimation.

The actual flow estimation procedure involved several steps. These steps are summarized below and described in detail in the following sections:

- Calculation of Flow Statistics - streamflow data was accessed from the STORET flow files for all streamflow gages previously "tied" to reaches. Average annual and monthly flows were derived directly from the STORET data base through the use of a computational program while low flow estimates and the degree of regulation of gages were determined through graphical techniques.

- Estimation of Flow Yields - flow yields related to geomorphologic/climatic characteristics (cfs/mi^2) were developed for each of the 2,111 USGS Cataloging Units in the contiguous United States. These yields were estimated separately for both average annual and low flow conditions.

- Selection of Gages - a set of criteria was established to select gages to be used in both area and flow estimating procedures. The criteria were designed to eliminate gages likely to produce erroneous streamflow estimates because of incorrect reach assignments or missing/incorrect data.

- Calculation of Area - data for selected gages were used to calculate the drainage area to stream length ratios (mi^2/mi) at gages. These ratios were used to estimate the area tributary to each reach and the total upstream drainage area. These gage data were supplemented by known areas for each USGS Cataloging Unit.

- Flow Estimation - the results of the analyses described above served as the basis for the actual streamflow estimating procedure. "Nominal" flow additions (average annual and low flow) for each reach were calculated from the area tributary to each reach and the appropriate flow yield derived from the geomorphologic/climatic characteristics of the reach. For those reaches where a flow gage was located on the reach, downstream of the reach, or the reach was bracketed by gages upstream and downstream of the reach, the nominal flow additions were linearly adjusted such that flow estimates matched streamflows at the gages and flow increased or decreased continuously between the gaged reaches. Special procedures were developed for estimating streamflows in unusual circumstances such as streams where flow decreased, streams entering from Canada, regulated streams, etc. Average monthly flows were calculated for each reach by proportioning the average annual flows on the basis of the monthly distributions at the controlling gages.

- Final Storage of Flow Data - the average annual, average monthly, and low flow estimates developed during the course of the project were stored by reach in computer data files established for each Hydrologic Region.

SECTION II

DATA BASE ANALYSIS

INTRODUCTION

Before the actual streamflows could be estimated it was necessary to analyze several facets of the available data base and develop various analytical techniques for preparing the data base for streamflow estimating. These steps included:

- development of flow statistics at gages
- analysis of flow statistics for low flows
- estimation of flow yields
- selection of gages
- estimation of drainage area

The purpose of this section is to present the various data base analyses and analytical techniques.

GAGED FLOW STATISTICS

Development of Average and Low Flows

Actual daily streamflow data were accessed from the STORET files for all streamflow gages previously "tied" to reaches. The U. S. Environmental Protection Agency (USEPA) developed a computer program to generate average annual, average monthly, annual 7-consecutive day low flows, and other flow statistics from the raw data. Average annual flow values were subsequently used directly to generate average annual flow yields as described below. The average monthly flows at each gage were used to estimate the distributions of average monthly flows in each reach. The annual 7-consecutive day low flows at each gage were further analyzed to develop low flow estimates for each gaged reach as presented in the following.

Low Flow Analysis

An estimation of the 7 day-10 year low flow (hereafter referred to as low flow) at each gage was required in order to utilize the gaged information in assigning low flows to reaches. The approach selected to calculate these low flows at gages utilized an interactive graphics program for USEPA's TEXTRONIX 4014-1 terminal based on the Velz-Gumbel graphical technique (II-1). In this method, flow data are plotted on extremal or log-extremal probability scales and a straight line distribution is visually fit to the data. Data not fitting the straight line distribution is indicative of streamflow regulation or other temporal changes in the physical environment upstream of the gage. When regulation is apparent it is possible to subtract or add a constant flow representing the regulated portion of total flow and to replot the unregulated flow distribution. An example of the Velz-Gumbel graphical method is presented in Figure II-1. The interactive graphics program simply makes it possible to quickly plot the arithmetic-or log-extremal distribution, adjust for regulation, replot, establish the 7-day-10 year flow and store the results. An example of the actual computer generated display is illustrated in Figure II-2. Low flow estimates for approximately 10,000 gages were developed in one man-month using this computerized approach.

ESTIMATION OF FLOW YIELDS

The procedure developed for estimating flows for each reach required the development of nominal average annual and low flow yields (cfs/mi^2) at gages assigned to reaches. The required yields were established from the analysis of measured flows for 9127 gaging stations.

It was perceived that flow yield is related to various geomorphologic/climatic characteristics of the drainage basin and that "partitioning" of flow yields by sets of these characteristics would yield a normal or log-normal distribution from which a mean value could be determined. This means a flow yield could then be assigned to all reaches manifesting the same set of geomorphologic/climatic characteristics. In this manner it was possible to develop a nation-wide, consistent set of flow yields for all reaches.

The geomorphologic/climatic characteristics selected for the analysis included:

- Mean annual precipitation
- Mean annual pan evaporation
- Mean annual high temperature frequency
- Geology
- Coefficient of variation of runoff
- Mean annual snowfall
- Land surface forms
- Topographic relief
- Average annual runoff
- Soils

In order to assure national consistency, USGS maps (II-2) of these characteristics were used to assign values of each characteristic to each USGS Catolog Unit and to each gage in the Reach File.

An interactive graphical procedure was used to display the probability distribution of flow yields in the actual "partitioning" process. Gaged data were selected on the basis of the various combinations of values of the geomorphologic/climatic characteristics and drainage area and displayed on the screen. A structured iterative process was employed to search for the best set of partionings based on the plotted distributions. This analysis was performed separately and independently for average annual flow and low flow. An example of the graphical program output is illustrated in Figure II-3. The results of this analysis included consistent nominal average annual and low flow yields for combinations of the spatial characteristics and drainage areas.

GAGE SELECTION

The gage information stored in the Analytical Gage File was used in the calculation of both drainage area and flow. A series of criteria were established to determine which gages were appropriate for use in either of these two calculations. These criteria were established in order to eliminate gages with insufficient data or those with a high probability of reach misassignment. The selection criteria used included:

- The "Integrity Code" supplied by EPA for each gage had to be "1". This indicates that the gage was identified by EPA as being on the assigned reach rather than on a tributary to that reach.

- A drainage area had to be assigned to the gage.

- The ratio of upstream miles of streams to drainage area had to be in approximately the same range as the average ratio for the Catalog Unit in which the gage was contained.

- The drainage area for a gage had to be greater than the sum of drainage areas for all gages immediately upstream.

- If a gage was to be used for flow estimation then both average and low flow information had to be available.

- A maximum of only one gage per reach was utilized. Preference was given to USGS gages and/or to the most downstream gage within a reach.

AREA ESTIMATION

The drainage area directly tributary to a reach and the total upstream drainage area for each reach were required prior to calculating flow estimates for each reach. The method of calculating tributary area for each reach was dependent upon the juxtaposition of the reach relative to gages that were selected as appropriate for area estimation. For those cases where the reach was upstream of a gage or located between gages, the average area per mile of stream for the area bracketed by the gages was utilized in conjunction with the reach length to calculate the tributary area. If there was no gage located downstream of the reach, then the average area per mile of stream calculated for the Catalog Unit in which the reach was contained was used. The former case was used in the calculation of the overwhelming majority of areas.

Total drainage area upstream of the downstream end of the reach was calculated by simply proceeding in an upstream to downstream direction and summing upstream and tributary areas. The one exception to this rule was the case where streams entered from Canada. Because areas were not available at the boundary, total drainage areas were not calculated on such streams until the first selected gage was encountered.

Application of the area calculation procedure is shown graphically in Figure II-4. For area A in Figure II-4, the area per mile ratio is calculated as the area at Gage 1 divided by the miles of stream in A. For area B, the area is calculated as the area at Gage 3 minus the area at Gages 1 and 2. For areas C and D, average area per mile of stream for the catalog units, in which the reaches are contained, are utilized. Total drainage area for all reaches below Gage 2 are calculated by the previously described procedure. Above Gage 2, total area is not calculated because the contributing drainage area in Canada is unknown.

REFERENCES

II-1. Velz, Clarence J., Applied Stream Sanitation, Wiley-Interscience, New York, 1970.

II-2. "The National Atlas of the United States of America," U. S. Department of the Interior, Geological Survey, Washington, D. C., 1970.

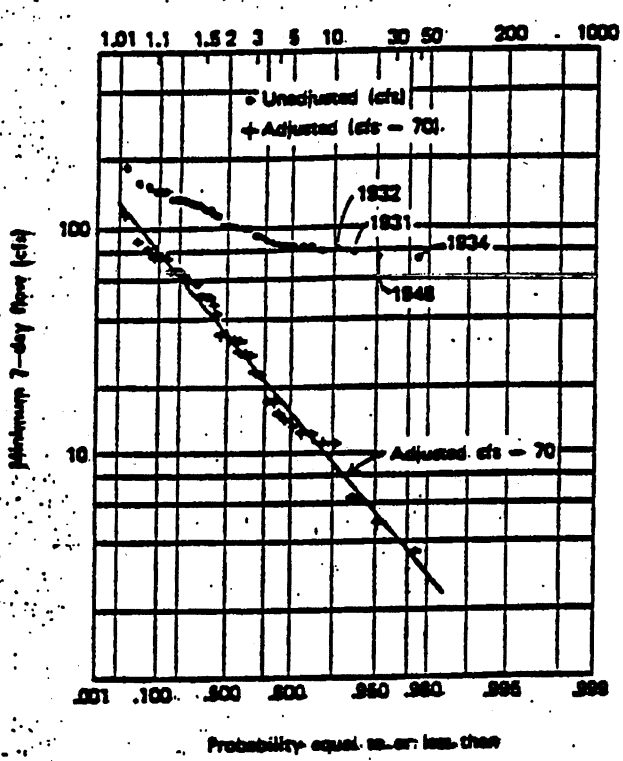
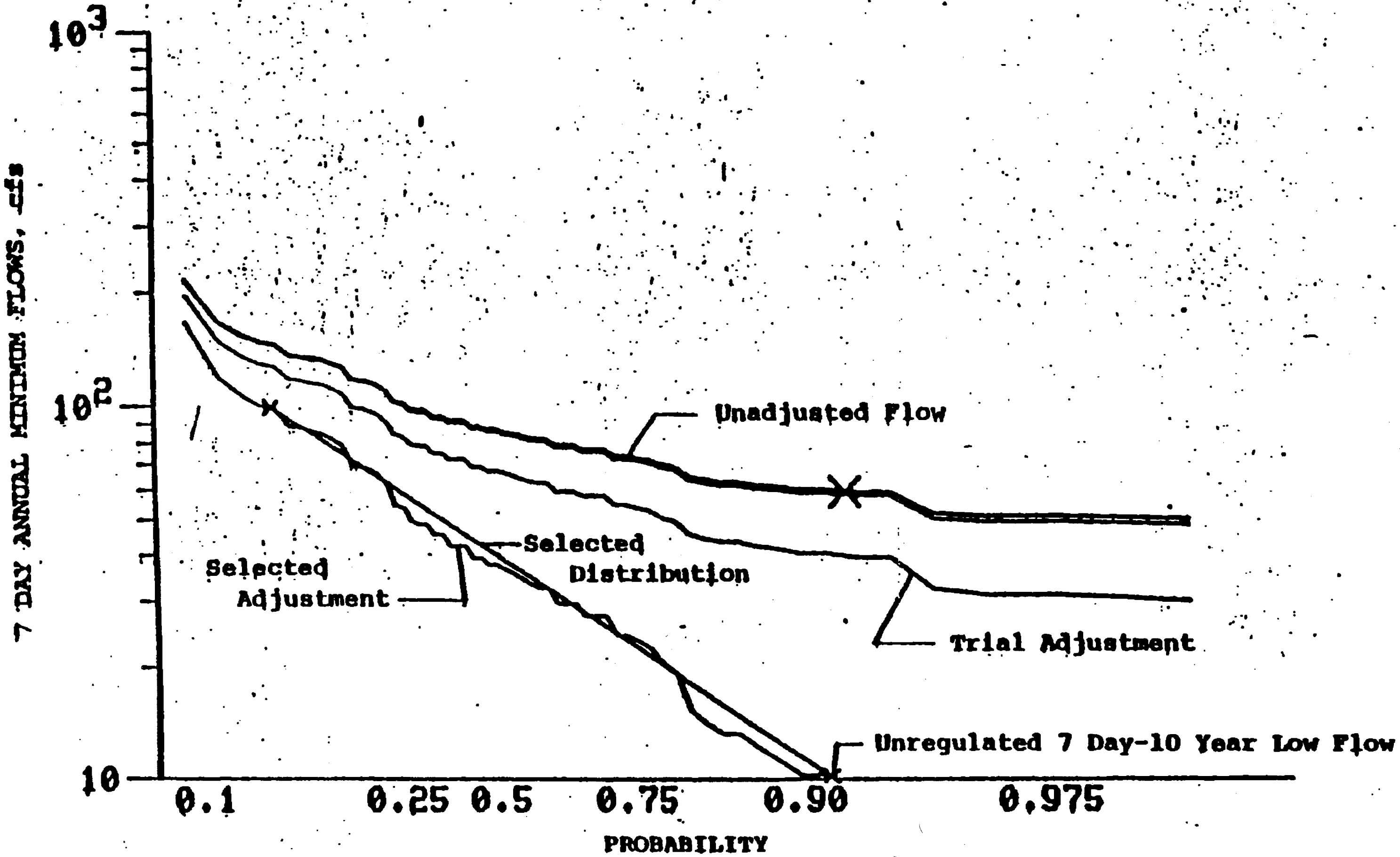


FIGURE II-1

EXAMPLE OF APPLICATION OF GUMBEL-VELZ GRAPHICAL METHOD
Reference (II-1)

FIGURE II-2

DISPLAY FROM THE INTERACTIVE GRAPHICS
LOW FLOW ANALYSIS PROGRAM



GAGE 04286000 SEQN # 13
MENU SELECTION?A
CODE ?? 1-91
INFO STORED: 7/10= 58.31 FLOW ADJ: -48.75 CODE 1
NO. OF YEARS OF RECORD 62 % ZERO 0.0

FIGURE II-3

DISPLAY FROM THE INTERACTIVE GRAPHICS PARTITIONING PROGRAM

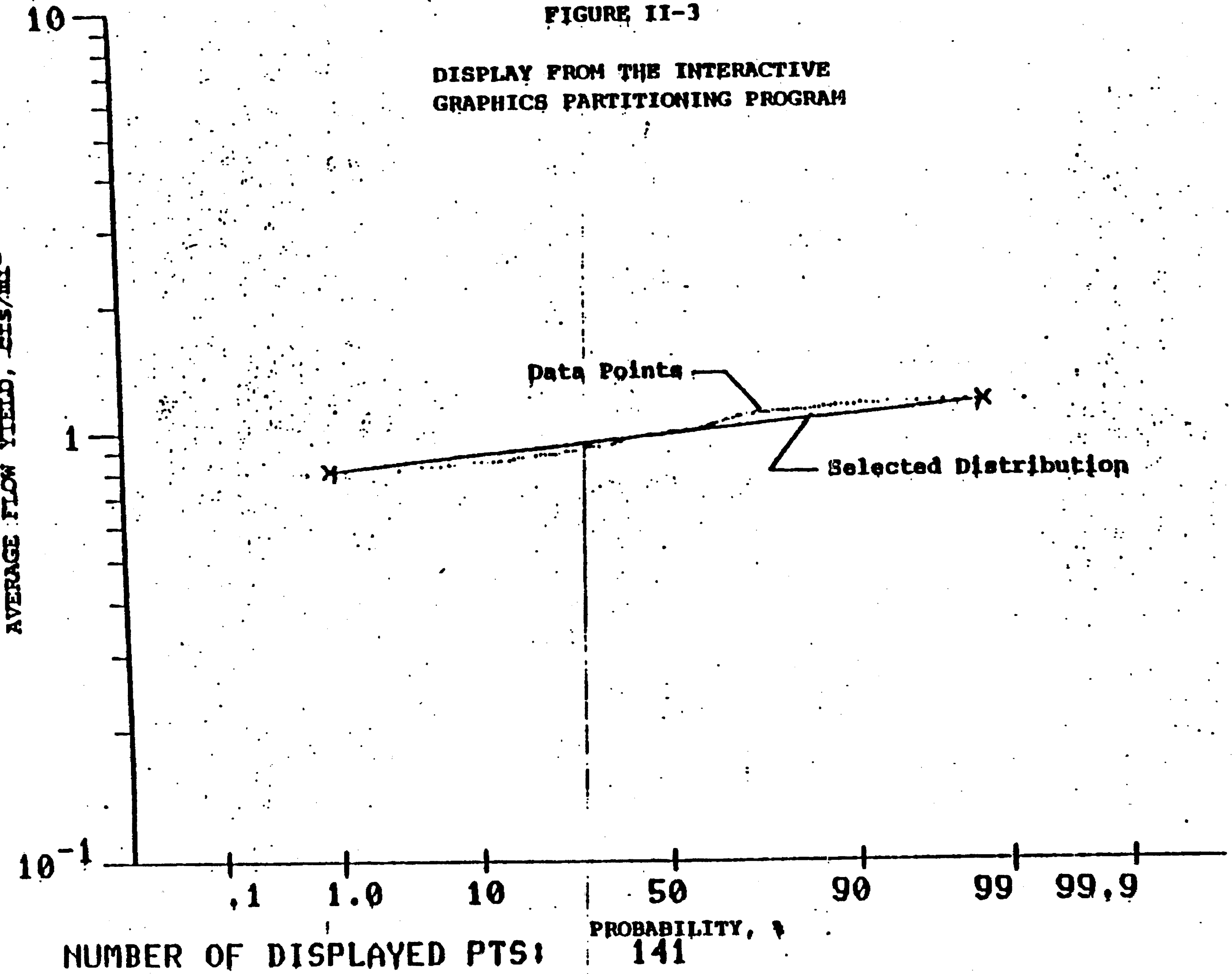
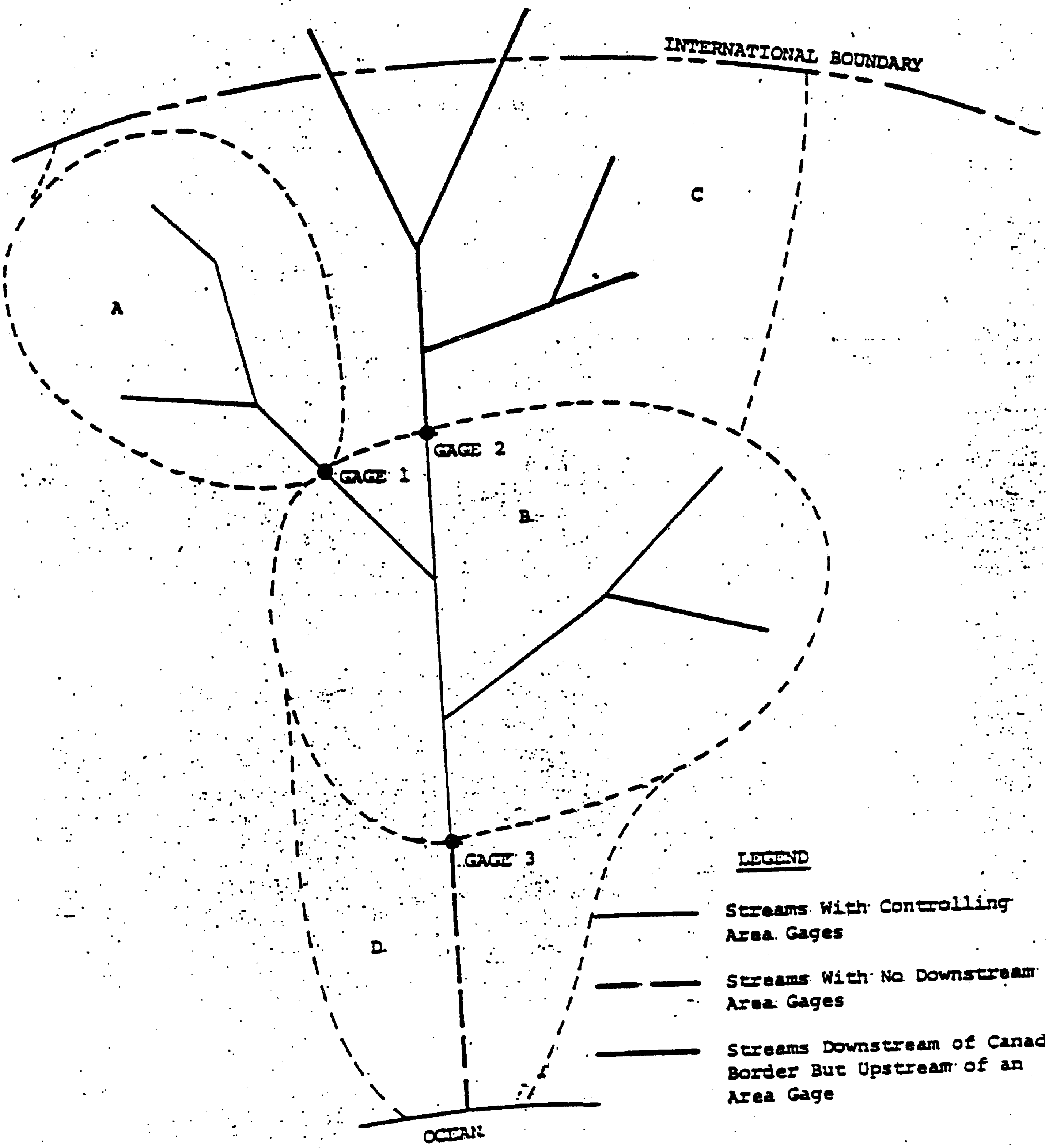


FIGURE II-4

AREA ESTIMATION PROCEDURES



SECTION III

FLOW ESTIMATION

INTRODUCTION

The actual flow estimation process was an integration of all of the specific data analyses and analytical techniques presented in Section II, organized within a consistent framework to permit the orderly calculations of the required flow data. The purpose of this section is to present the details of this process.

FLOW ESTIMATING TECHNIQUES

General Approach

In general, final independent streamflow estimates for average annual and low flow were developed in hydrologic order, upstream to downstream; calculating lateral flow contributions and adding these to upstream contributions as one moved downstream. This process was followed until flows for an entire Hydrologic Region were computed. An "interbasin" file was established to facilitate flow transfers from region to region.

The analytical technique for estimating streamflow can best be described by example, as illustrated in Figure III-1. Nominal flow yields for each reach were determined independently for average annual and low flow conditions within each Cataloging Unit based on the values of the geomorphologic/climatic characteristics within the Unit containing the reach and the total drainage area upstream of the reach. Nominal flow increments were calculated for each reach by multiplying the nominal flow yields by area and summing these flow increments for all reaches between a set of gages. As illustrated in Figure III-1, this process was performed for the four reaches between Gage A and Gage B. The sums of the nominal flows were then compared to the incremental gage flow (measured flow of Gage B minus measured flow at Gage A) and an adjustment factor calculated as the ratio between the gaged incremental and nominal incremental flows. This factor was then applied to the nominal flow increment for each reach to determine the adjusted incremental flows.

This analytical technique was used to estimate streamflows for reaches in areas with known gage flow increments such as Areas D, E, and F in Figure III-2. However, the Reach File contained many reaches for which there were no associated gage flow increments. Examples of these conditions are illustrated in Figure III-2; such as:

- Area A where streams enter from Canada, thus there is no upstream gage in the File;
- Area B where there is no gage on a terminal set of reaches; and
- Area C which contains reaches with no downstream gages.

Flow estimates for these types of areas in the Reach File were based solely on the nominal flow increments without gage adjustments.

Modified Techniques

Modifications in the general flow estimating approach were necessary in a few instances where:

- stream diversions exist
- streams lose flow (and no diversions are apparent)
- streams enter from Canada
- streams are regulated by dams and other man made structures

Stream diversions, i.e. streams which split into two outflows, are represented in a specific manner in the Reach File. Within the File, one of the outflows is designated the major outflow and the other as the minor outflow. Two options were developed for estimating flow in both outflows. In the first case the actual distribution of flow is presumed to be known (or at least approximated) and is specified in the Reach File. Under this option a percentage was used to assign the total flow to the two outflow reaches (e.g. 70 percent of the flow in the major and 30 percent in the minor). If the flow distribution was not known a second option for estimating flow was exercised which simply assigned all of the flow to the major outflow.

For those cases where the incremental gaged flow was negative (i.e. the streamflow decreases in the downstream direction) the following methodology was applied. Along the flow path between gages, an exponential loss function was used with the exponential decay constant ($miles^{-1}$) selected such that a smooth function between flow gages resulted. The exponential function model assumes that the rate of loss is always proportional to the flow at a point. This function guarantees both a smooth continuous loss rate and no negative flows. Tributaries to reaches losing flow were assumed to have zero flows. An example of the application of the exponential loss function is presented in Figure III-3.

For those streams entering from Canada, the flows entering at the border were not known. For all reaches downstream of the border but upstream of a flow gage, the incremental nominal flows were calculated. If the total incremental flow exceeded the gage flow then the incremental flows were adjusted and it was assumed that no flow entered from Canada. If the gage flow exceeded the total incremental flow, the difference was assumed to originate in Canada and was assigned to the stream entering from Canada if there was only one stream or to the largest stream entering from Canada if there were several streams. The largest stream was determined from the "stream level" a stream descriptor available in the Reach File. A level one stream is one that is terminal (e.g. discharges to an ocean); a level two stream is tributary to a level one stream, etc. For example, in the network shown in Figure III-4, if the flow at Gage G was 1000 cfs and the nominal incremental flow above it was 300 cfs then the flow entering the level one stream at point A was estimated to be 700 cfs while the flow entering at the higher level streams at points B, C and D were assigned a stream flow of zero.

A certain number of gages were determined to manifest controlled or regulated low flows during the low flow analyses (Section II). Incremental streamflow between pairs of regulated gages or from a regulated gage downstream to an unregulated gage required special attention because the regulated portion of flow behaves differently than does the unregulated portion. In Figure III-5 Gage A and Gage B were both determined to be regulated. Total streamflow between the gages was estimated by first computing the unregulated flow using the general approach for estimating flow presented earlier and then adding the interpolated portion of regulated flow between gages to each reach between the gages.

Summary

The general approach and the various modifications required by the structure of the Reach File were used to generate average annual and low flow estimates for all reaches in the Reach File. Mean monthly flows were calculated by applying the ratio of each monthly average flow to annual mean flow associated with the first gage downstream of each reach to the computed average annual flow for the reach. For reaches with no downstream gages (see Figure III-2) each average monthly ratio of all gages in the Catalog Unit was used as the surrogage.

OUTPUT FILES

The final estimated flow data were stored in a sequential file generated for each Hydrologic Region. This file is numerically ordered by reach number and in addition to containing the flow data the file also contains a four digit code describing the characteristics of the flow estimating technique used. Starting at the left-most digit the codes are:

- Digit 1: Is reach between the Canadian border and a stream gage (1 = Yes, Blank = No)

- Digit 2: Relationship between reach and stream gage
  2 - There is a gage on the reach
  1 - There is a gage downstream of the reach
  0 or blank - There is no downstream gage

- Digit 3: Relationship between calculated average and low flows
  1 - Low flow exceeds average flow
  0 or blank - Average flow greater than or equal to low flow

- Digit 4: Not used - always set to zero

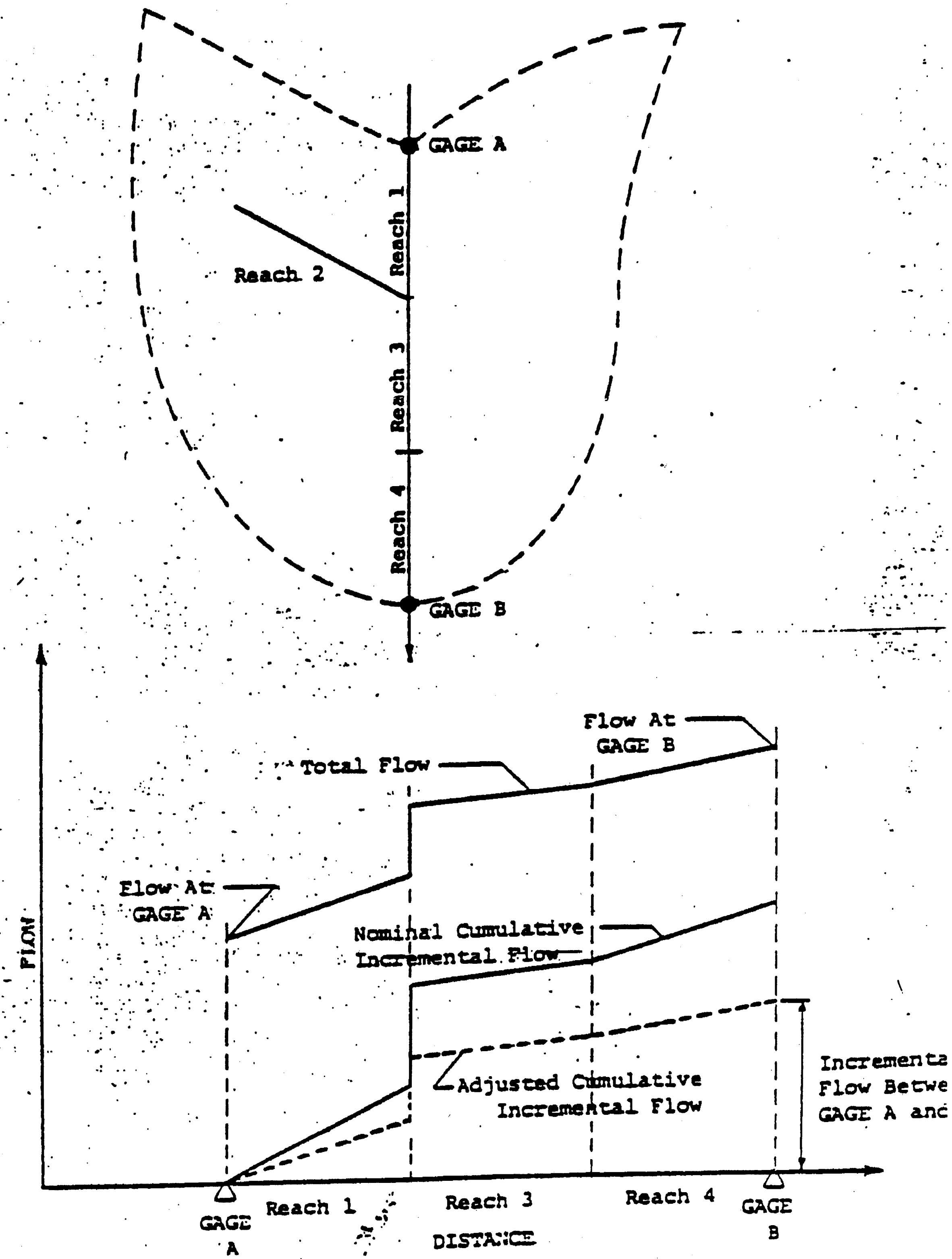
The 18 regional output files are stored as sequential, unformatted files. An example of the file structure is presented in Figure III-6.

Lest the reader feels that a "1" in the third digit (low flow exceeds average annual flow) means that either a mistake has been made or the flows are not to be trusted or, worse, that such an anomaly means no flow data in the file is useable, it is important for the reader/user to understand how such an anomaly was created. Such a situation was created when the incremental low flow between sequential gages exceeds the incremental average annual flow between the paired gages. Under these circumstances the incremental low flow yield (cfs/mi^2) exceeds the incremental average annual flow yield (cfs/mi^2), thus any tributary reach between the gages with no incoming flow was assigned a low flow greater than the average annual flow. In the example in Figure III-7 the incremental low flow between Gage A and Gage B is 150 cfs and exceeds the incremental average annual flow of 100 cfs. Because of the upstream inflow at Gage A, the low flow will not exceed the average annual flow in the mainstem reaches between Gage A and Gage B, but the tributary reach (Reach 2), having no upstream inflow, is assigned a low flow in excess of the average annual flow.

This situation is relatively rare in the File but was caused by gages incorrectly located, errors in the available flow statistics, or a significant water withdrawal which impacts average annual flow but not low flow. As noted, these situations are "flagged" in the flow file; however, in order to maintain consistancy in the methodology and the lack of any supplementary information for determining the cause of the anomoly, no adjustments were made in these flow estimates.

FIGURE III-1

GENERAL FLOW ESTIMATING APPROACH



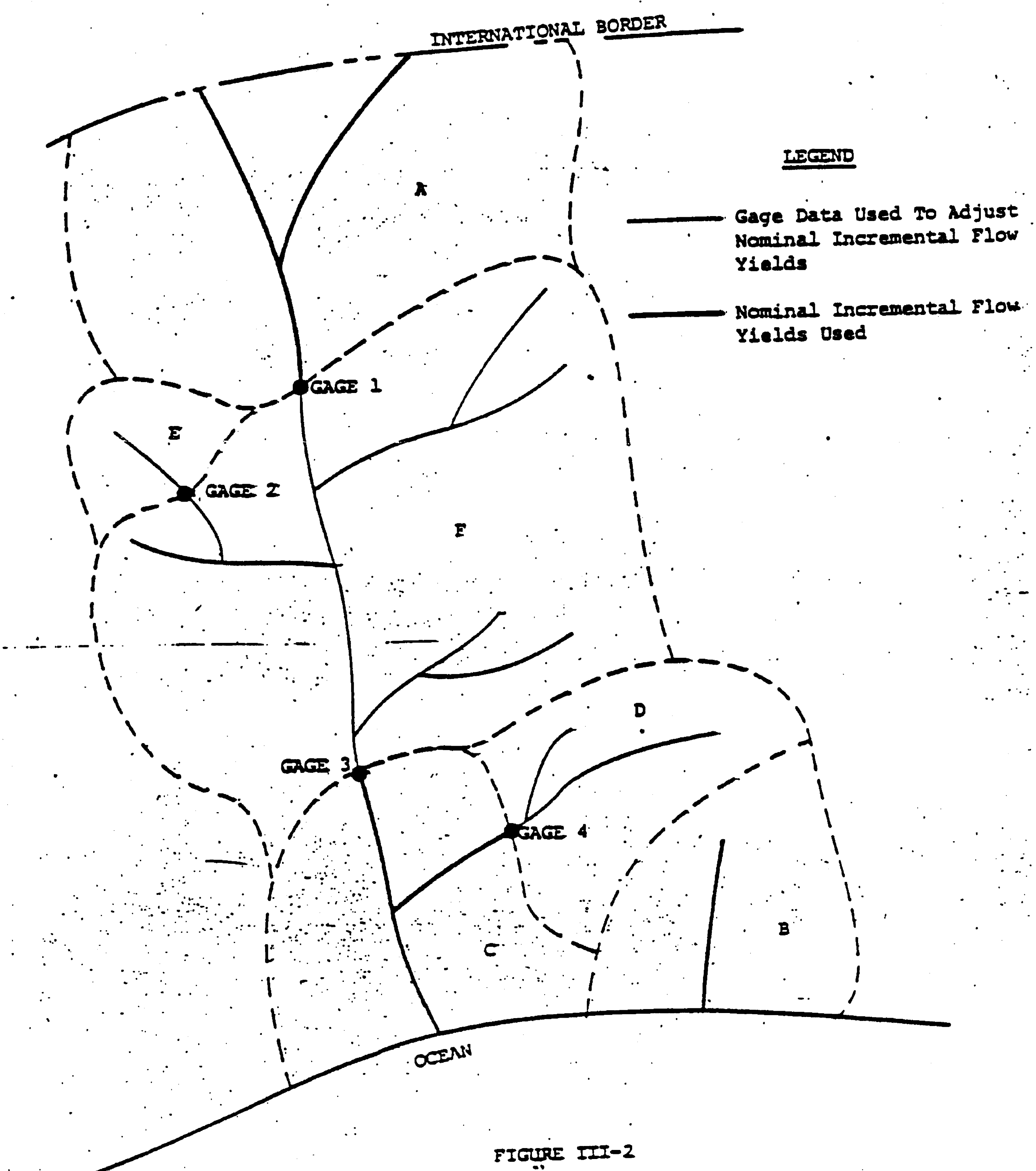


FIGURE III-2

STREAM CONFIGURATION ASSOCIATED WITH FLOW ESTIMATION METHODS

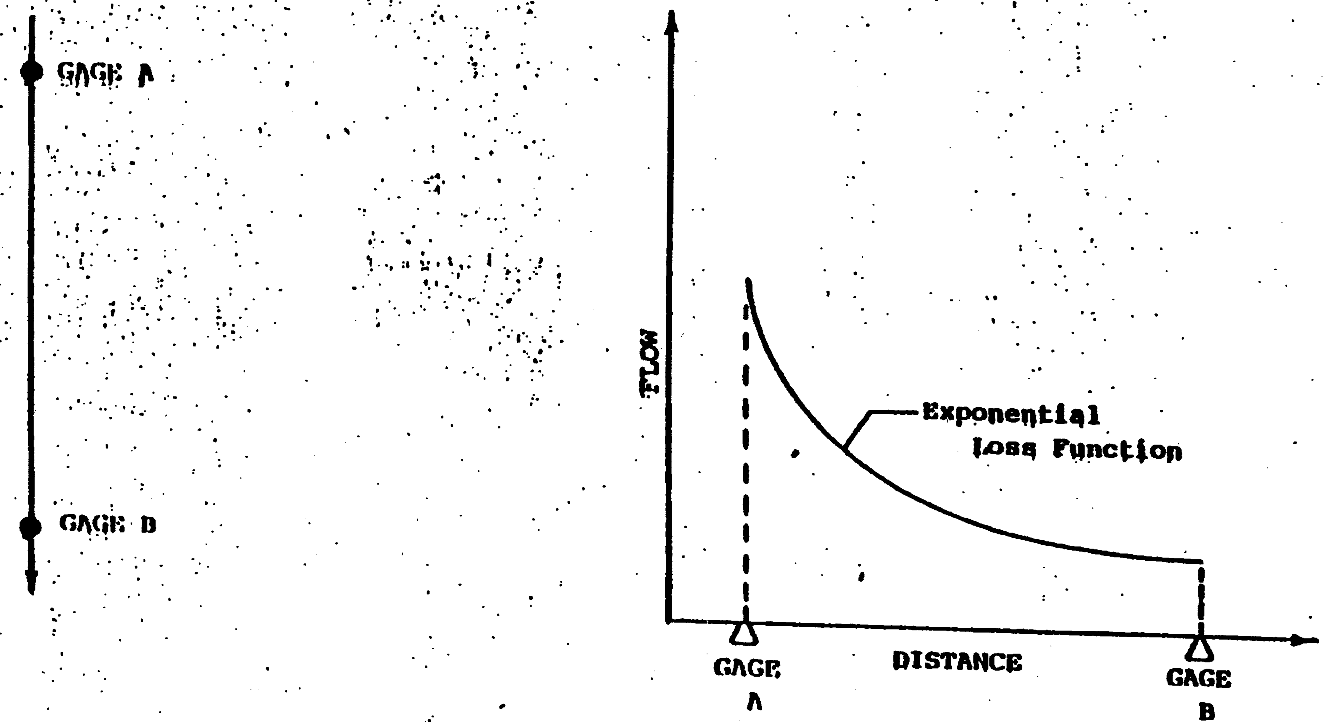


FIGURE III-3

EXPONENTIAL LOSS FUNCTION

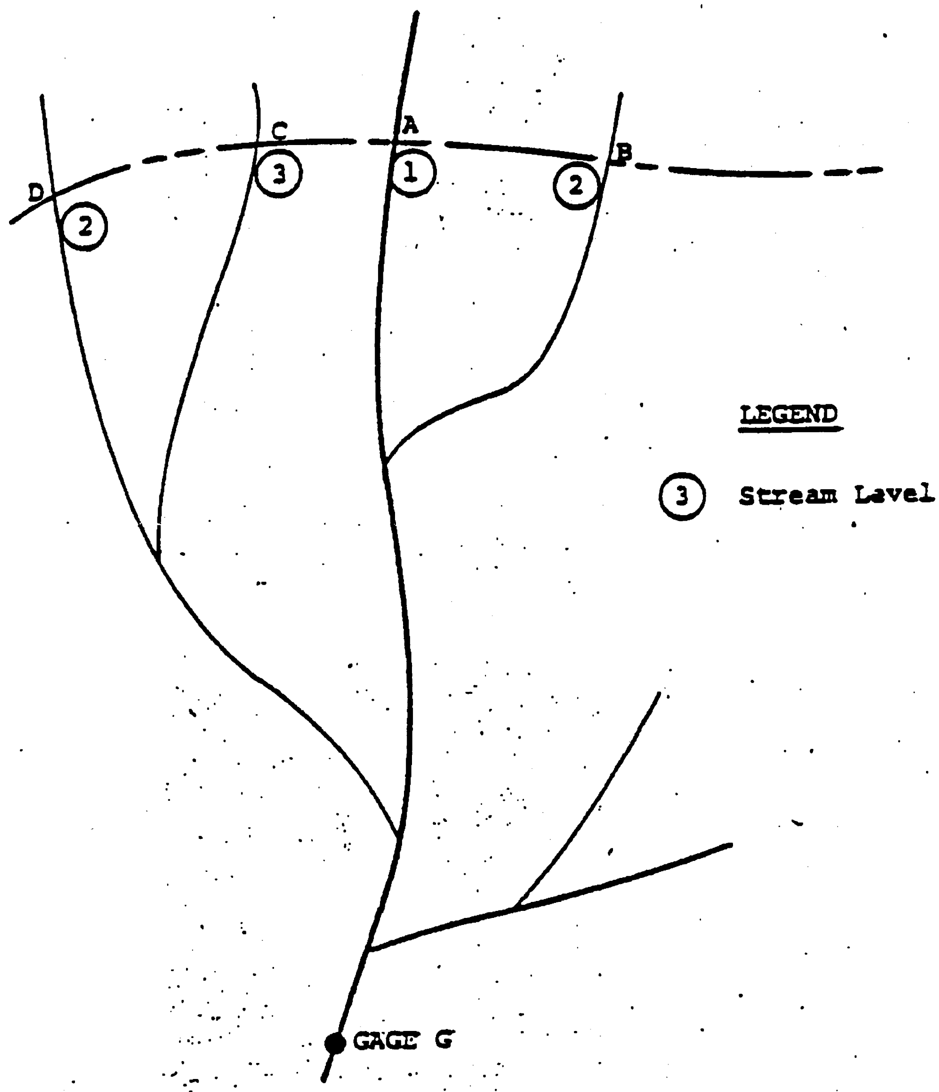


FIGURE III-4

STREAM NETWORK ENTERING FROM CANADA

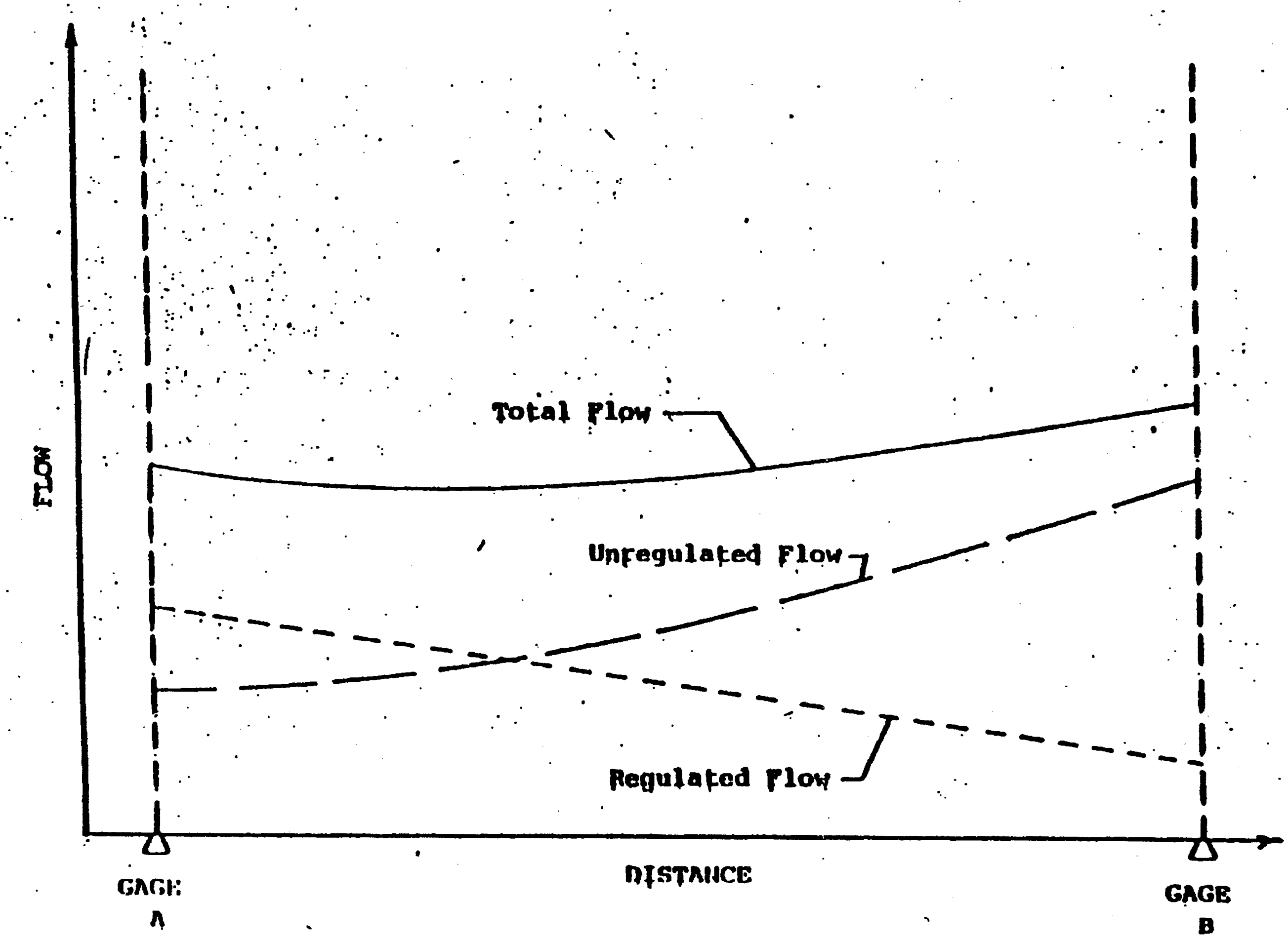


FIGURE III-5

EXAMPLE OF FLOW ESTIMATION FOR REGULATED STREAMS

FIGURE III-6

EXAMPLE LISTING OF FLOW FILES

BEGIN PROCESS FOR REGION   V FILE CREATED ON 820326 AT 93120

CU	SEG	AVG F.	LOW F.	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC	CODE
Y 10001	1	151.	0.	14.	15.	27.	447.	582.	357.	173.	70.	34.	33.	25.	16.	100
Y 10001	2	135.	0.	14.	13.	24.	401.	523.	320.	155.	62.	31.	30.	23.	15.	100
Y 10001	3	120.	0.	11.	12.	21.	354.	462.	283.	137.	55.	27.	26.	20.	13.	100
Y 10001	4	11.	0.	1.	1.	2.	34.	44.	27.	13.	5.	3.	2.	2.	1.	100
Y 10001	5	98.	0.	9.	10.	17.	291.	379.	232.	112.	45.	22.	21.	16.	11.	100
Y 10001	6	48.	0.	5.	5.	9.	143.	186.	114.	55.	22.	11.	11.	8.	5.	100
Y 10001	7	37.	0.	4.	4.	7.	109.	142.	87.	42.	17.	8.	8.	6.	4.	100
Y 10001	10	24.	0.	2.	2.	4.	68.	89.	55.	26.	11.	5.	5.	4.	3.	100
Y 10001	11	9.	0.	1.	1.	2.	27.	35.	21.	10.	4.	2.	2.	1.	1.	100
Y 10001	12	15.	0.	1.	1.	3.	44.	58.	35.	17.	7.	4.	3.	3.	2.	100
Y 10001	14	190.	0.	1.	1.	179.	568.	186.	107.	103.	13.	14.	2.	2.	1.	200
Y 10001	16	110.	0.	1.	1.	198.	627.	205.	121.	113.	14.	14.	3.	2.	1.	0
Y 10001	18	197.	0.	2.	2.	353.	1120.	366.	215.	202.	25.	25.	5.	3.	2.	0
Y 10002	1	40.	0.	1.	2.	54.	175.	90.	57.	30.	19.	18.	14.	8.	3.	200
Y 10002	2	13.	0.	0.	1.	17.	55.	28.	19.	9.	6.	6.	4.	2.	1.	100
Y 10002	3	1.	0.	0.	0.	1.	2.	1.	1.	0.	0.	0.	0.	0.	0.	100
Y 10002	4	4.	0.	0.	0.	8.	27.	14.	9.	5.	3.	3.	2.	1.	1.	100
Y 10002	5	6.	0.	0.	0.	8.	25.	13.	9.	4.	3.	3.	2.	1.	0.	100
Y 10003	2	267.	0.	25.	26.	47.	789.	1029.	630.	305.	123.	60.	58.	44.	27.	200
Y 10003	3	2.	0.	0.	0.	0.	6.	8.	5.	2.	1.	0.	0.	0.	0.	100
Y 10003	4	262.	0.	25.	26.	46.	775.	1010.	618.	299.	121.	59.	57.	44.	28.	100
Y 10003	5	248.	0.	24.	24.	44.	735.	959.	587.	284.	114.	56.	54.	41.	27.	100
Y 10003	6	246.	0.	24.	24.	43.	729.	950.	582.	282.	113.	56.	54.	41.	27.	100
Y 10003	7	218.	0.	21.	21.	38.	644.	840.	515.	249.	109.	49.	47.	36.	24.	10
Y 10003	8	218.	0.	21.	21.	38.	644.	840.	514.	249.	100.	49.	47.	36.	24.	10
Y 10003	11	217.	0.	21.	21.	38.	644.	839.	514.	249.	100.	49.	47.	36.	24.	10
Y 10003	12	217.	0.	21.	21.	38.	643.	839.	514.	249.	109.	49.	47.	36.	24.	10

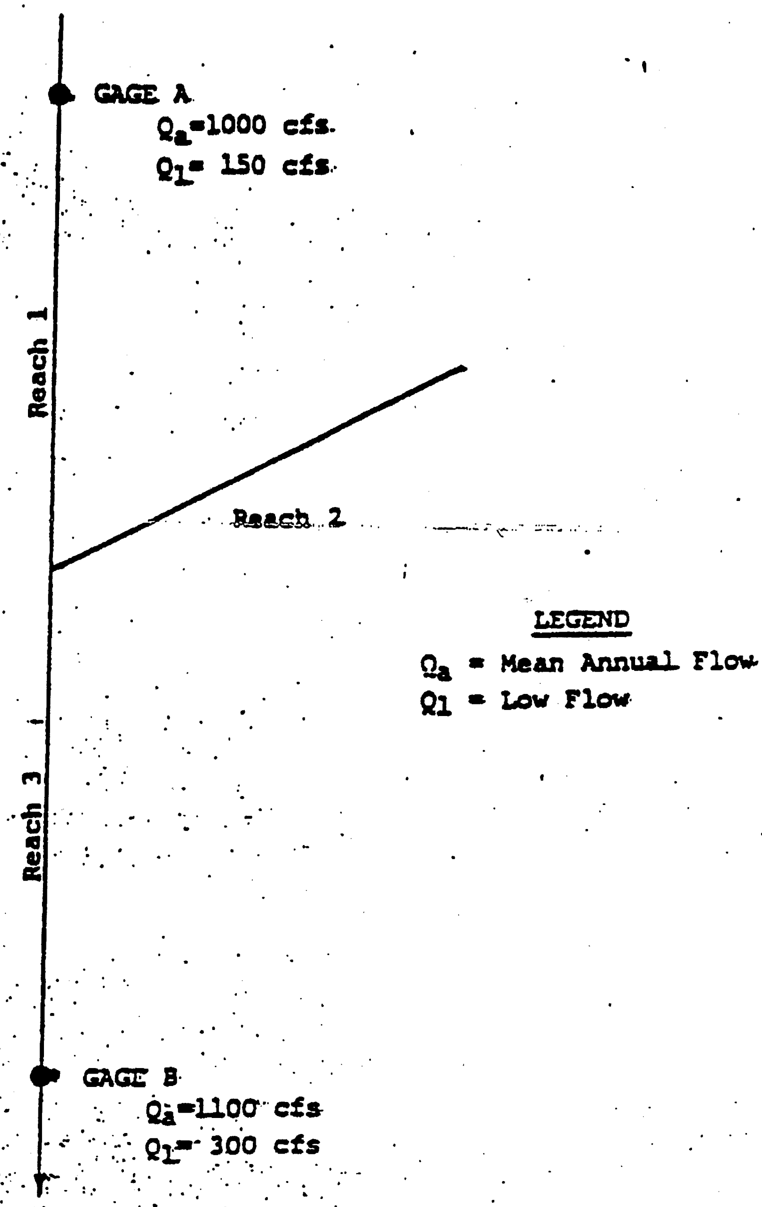


FIGURE III-7

EXAMPLE LEADING TO LOW FLOW
ESTIMATE EXCEEDING AVERAGE FLOW ON TRIBUTARY